Sensitivity Analysis in Testing the Consistency of Vulnerability Models

by Joseph C. Collins

20020319 205

Army Research Laboratory

Aberdeen Proving Ground, MD 21005-5068

ARL-TR-2668   February 2002

Sensitivity Analysis in Testing the Consistency of Vulnerability Models

Joseph C. Collins
Survivability/Lethality Analysis Directorate, ARL

Abstract

Comparison of series of live-fire test results with stochastic simulation model output has been a long-standing problem. The usual approach, the order by probability (OP) test, allows neither for quantification of typical adjustments made in the implementation of the test nor for sensitivity analysis. A new test is presented here which may be used in conjuction with the OP test to provide these missing features.

Acknowledgments

INTENTIONALLY LEFT BLANK.

Contents

INTENTIONALLY LEFT BLANK.

List of Figures

List of Tables

INTENTIONALLY LEFT BLANK.

1. Vulnerability Models

The task at hand is to compare a set of live-fire (LF) results as a group to stochastic simulation model output in order to determine if the model adequately predicts the LF results. The simulation is replaced by a simple probabilistic model, and statistical tests are conducted to determine if the LF results are consistent with the idealized simulation distribution.

Assume initially that a stochastic vulnerability/lethality (VL) model predicts either no loss of function (LoF) or complete LoF for a particular component, system, or vehicle. Under a fixed set of input conditions, output typically varies due to the stochastic nature of the model. The output of such a model can be represented as $X \in \{0,1\}$, where $\Pr[X = 1] = p = 1 - \Pr[X = 0]$ and the single parameter p of the model lies in the interval $[0,1]$. The values of X have the interpretation that $X = 1$ indicates complete LoF and $X = 0$ indicates no LoF, so that X is in fact the LoF. This random variable X has a Bernoulli distribution with parameter p. Equivalently, X is a binomial random variable with sample size 1 and parameter p, denoted $X \sim B(1,p)$.

The model independently exercised under k sets of input conditions (each representing a distinct shot with a particular target, threat, geometry, etc.) is denoted by the vector $X = (X_i)_{i=1}^{k} = (X_1, X_2, \ldots, X_k)$ with vector parameter $p = (p_i)_{i=1}^{k}$. Each $X_i \in \{0,1\}$ is the LoF for configuration i, and $\Pr[X_i = 1] = p_i = 1 - \Pr[X_i = 0]$. The distribution of the vector X is k-dimensional binomial of sample size 1 and parameter p with independent components, or $X \sim B_k(1,p)$. Component i of k has the binomial distribution $X_i \sim B(1,p_i)$.

Each set of conditions is associated with a single experimental result in $L_i \in \{0,1\}$, which together constitute the data vector $L = (L_i)_{i=1}^{k} \in S = \{0,1\}^k$. So L is a k-vector of zeros and ones, and the sample space has size 2^k.

As VL models increase in complexity, they allow for partial LoFs, so $X \in [0,1]$. For a simple example, suppose that a vehicle has a maximum speed of 60 mph. If the vehicle can maintain that speed after the shot, then $X = 0$ since there is no LoF. If the vehicle cannot move, then there is total LoF and $X = 1$. And if the vehicle can operate at a maximum speed of 30 mph, then one might assess the LoF at $X = 0.5$.

Typically the number of outputs for a single input configuration is finite and in fact completely exhausted after a few hundred or a few thousand iterations, giving perhaps as many as a dozen distinct output LoFs.

In one dimension, the model output LoF, denoted X, can take on any of the n values of $x = (x_j)_{j=1}^n$ with respective probabilities $p = (p_j)_{j=1}^n$, so that $\Pr[X = x_j] = p_j$ where $\sum_{j=1}^n p_j = 1$ and each $p_j > 0$. The one-dimensional model for X is finite discrete, or multinomial with sample size one, value parameter x, and probability parameter p. This is denoted $X \sim M(1, x, p)$, and the sample space or set of possible values of X is $\{x_j\}_{j=1}^n$.

Again, the model can be exercised under k sets of input conditions, each representing a distinct shot. The VL model is then k-dimensional vector multinomial $M_k(1, x, p)$ of size 1 where the value parameter list is $x = (x_i)_{i=1}^k$ with $x_i = (x_{i,j})_{j=1}^{n_i}$ and the probability parameter list is $p = (p_i)_{i=1}^k$ with $p_i = (p_{i,j})_{j=1}^{n_i}$ for each i of k, so that $\Pr[X_i = x_{i,j}] = p_{i,j}$. This means that for shot i, the LoF value $x_{i,j}$ occurs in the model with probability $p_{i,j}$. In this case, it is assumed that the data vector of LF LoF assessments $L = (L_i)_{i=1}^k$ lies in the sample space $S = \otimes_{i=1}^k \{x_{i,j}\}_{j=1}^{n_i}$, which means that each single result L_i lies in the appropriate set $\{x_{i,j}\}_{j=1}^{n_i}$.

A final generalization occurs when L does not necessarily lie in S. For example, it may be reasonable to conclude that in one dimension the LF LoF assessment $L = 0.26$ is consistent with the model $X \sim M_k(1, x, p)$ where LoF values $x = (0.00, 0.25, 0.50)$ have respective probabilities $p = (0.10, 0.85, 0.05)$, whereas $L = 0.99$ is not. In this case, it is desirable to test whether L is close to X in some way.

The tests under consideration here assume that L is an $M_k(1, x, \theta)$ vector, perhaps $B_k(1, \theta)$, or close to an $M_k(1, x, \theta)$ vector. Furthermore, the VL model is assumed to be $X \sim M_k(1, x, p)$. These tests all take the form of either

$$H_0 : \theta = p$$
$$H_1 : \theta \neq p \tag{1}$$

or

$$H_0 : \mathrm{E}\, L = \mathrm{E}\, X$$
$$H_1 : \mathrm{E}\, L \neq \mathrm{E}\, X. \tag{2}$$

The model parameter (x, p) is estimated from n iterations of the VL model, so the output is $M_k(n, x, p)$. It is thus possible to provide error bounds, etc., for the estimates of (x, p). This estimation problem is ignored without further mention, and (x, p) is taken to be the true parameter of the VL model.

More detailed information about the concepts of probability and statistics used here are presented in the Appendix.

2. The Order by Probability (OP) Test

The OP test was originally designed to check the consistency of component-wise or system-wise VL model predictions for a single threat-target interaction, where each vector dimension represents a single component or system in the vehicle.[1] Each of k components survives or is destroyed independently, so the appropriate model for a single shot against a single vehicle is vector binomial $B_k(1, \theta)$. The parameter vector is $\theta = (\theta_i)_{i=1}^{k}$, in which θ_i represents the probability that the ith component is killed. A modification of the original test exploits the empirical distribution of the simulation vector for use when the component or system kill probabilities are deemed nonindependent. A reliable estimate of the empirical simulation distribution may be difficult to obtain. For example, a set of 20 components presents over one million possible outcomes; furthermore, a large sample may be required to accurately characterize the relative frequency of occurrence of the less likely outcomes.

However, distinct threat-target interactions (separate shots) are independent, so the binomial/multinomial OP test is applicable to a set of separate-shot complete-vehicle LoFs. The original formulation of the OP test applies when the response and predictions are in $\{0, 1\}$ for each dimension. The test can be extended to the case of the $M_k(1, x, \theta)$ version of the VL model discussed in the previous section, where vehicle LoFs lie in $[0, 1]$. The ith threat-target interaction of k assumes each of its n_i possible LoF values $(x_{i,j})_{j=1}^{n_i}$ with respective probability $p_{i,j}$, and the interactions are independent. So the simulation distribution is taken to be vector multinomial of sample size 1.

A technical description of the OP test follows, first in the original binomial form and then in the multinomial form. This level of representation engenders formulae that easily translate into computational instructions.

2.1 Technical Description

In the binomial case the simulation distribution is $X \sim B_k(1, p)$, and the experiment distribution is assumed to be $L \sim B_k(1, \theta)$. The OP hypothesis test is

$$H_0 : \theta = p$$
$$H_1 : \theta \neq p. \tag{3}$$

---

[1] Webb, D. W. "Tests for Consistency of Vulnerability Models." BRL-TR-3030, U.S. Army Ballistic Research Laboratory, Aberdeen Proving Ground, MD, August 1989.

Since $\mathrm{E}\,Z = \theta$ when $Z \sim B_k(1,\theta)$, this is the same as

$$H_0 : \mathrm{E}\,L = \mathrm{E}\,X$$
$$H_1 : \mathrm{E}\,L \neq \mathrm{E}\,X. \tag{4}$$

The OP test is a generalized likelihood ratio test. The sample distribution is $Z \sim B_k(1,\theta)$, where the parameter space is $\Theta = [0,1]^k$ and the null space is $\Theta_0 = \{p\}$. The test of

$$H_0 : \theta \in \Theta_0$$
$$H_1 : \theta \in \Theta - \Theta_0 \tag{5}$$

uses the generalized likelihood ratio $\lambda(L)$ as a test statistic. The generalized likelihood ratio function is

$$\lambda(t) = \frac{\sup_{\theta \in \Theta_0} \mathrm{Pr}_\theta[Z = t]}{\sup_{\theta \in \Theta} \mathrm{Pr}_\theta[Z = t]} = \mathrm{Pr}_p[Z = t] = f_p(t), \tag{6}$$

where f_p is given by equation (A-9) in the Appendix, so the test statistic is $\lambda(Z) = f_p(Z)$. Construction of the test statistic distribution follows. Let $\{\tau_n\}_{n=1}^{2^k}$ be an enumeration of the sample space $S = \{0,1\}^k$ so that $f_p(Z)$ takes the values $\{f_p(\tau_n)\}_{n=1}^{2^k}$. Then the distribution of $f_p(Z)$ is given by

$$\mathrm{Pr}_\theta[f_p(Z) = f_p(t)] \;\; = \sum_{f_p(\tau_n)=f_p(t)} \mathrm{Pr}_\theta[Z = \tau_n] \;\; = \sum_{f_p(\tau_n)=f_p(t)} f_\theta(\tau_n), \tag{7}$$

where, again, f_θ is as in equation (A-9). Now, $f_p(Z)$ has small values significant. The order can be reversed by using for a test statistic either $T(Z) = g_p(Z) \equiv 1 - f_p(Z)$, or, equivalently, the rank $T(Z) = g_p^*(Z)$ of $g_p(Z)$ in $\{g_p(\tau_n)\}_{n=1}^{2^k}$, both of which have large values significant. Note that $g_p^* = f_p^-$, the descending rank of $f_p(Z)$ in $\{f_p(\tau_n)\}_{n=1}^{2^k}$.

In either case, the cdf of $T(Z)$ is

$$F_\theta(t) = \sum_{n=1}^{2^k} f_\theta(\tau_n) I_{[T(\tau_n),\infty)}(t), \quad \text{i.e.,}$$
$$= \sum_{n=1}^{2^k} f_\theta(\tau_n) I_{[g_p(\tau_n),\infty)}(t) \quad \text{or}$$
$$= \sum_{n=1}^{2^k} f_\theta(\tau_n) I_{[g_p^*(\tau_n),\infty)}(t), \tag{8}$$

4

and the null distribution is given when $\theta = p$. The decision rule is to reject if $T(Z) > T_\alpha$, where the critical value is the usual $T_\alpha = Q_p(1 - \alpha_o)$. The power of the test under the alternative θ is $1 - \beta = 1 - F_\theta(T_\alpha)$.

Extension to the multinomial VL model is transparent. In this case, the sample distribution is $Z \sim M_k(1, x, \theta)$. The hypothesis test again takes the form

$$H_0 : \theta = p$$
$$H_1 : \theta \neq p, \tag{9}$$

where the simulation distribution is $X \sim M_k(1, x, p)$ and the experiment distribution is $L \sim M_k(1, x, \theta)$. Note, the values x that the model and data take are fixed and assumed to be identical. It is the probability of occurrence θ of these values that is being tested. In terms of components, the parameter of interest is $\theta = ((\theta_{i,j})_{j=1}^{n_i})_{i=1}^k$. In the multinomial case, the test is not a test of $\mathrm{E}\,L = \mathrm{E}\,X$, since $\mathrm{E}\,Z = (\sum_{j=1}^{n_i} \theta_{i,j} z_{i,j})_{i=1}^k$ for $Z \sim M_k(1, z, \theta)$. At any rate, this is another generalized likelihood ratio test, and construction of the test statistic is analogous to the binomial case. The generalized likelihood ratio is

$$\lambda(t) = \frac{\sup_{\theta \in \Theta_0} \mathrm{Pr}_{x,\theta}[Z = t]}{\sup_{\theta \in \Theta} \mathrm{Pr}_{x,\theta}[Z = t]} = \mathrm{Pr}_{x,p}[Z = t] = f_{x,p}(t), \tag{10}$$

where $f_{x,p}$ is given by equation (A-13), so the test statistic is $f_{x,p}(Z)$. Let $\{\tau_n\}_{n=1}^N$ be an enumeration of the sample space $S = \otimes_{i=1}^k \{x_{i,j}\}_{j=1}^{n_i}$, where $N = \prod_{i=1}^k n_i$. Then $f_{x,p}(Z)$ is defined on $\{f_{x,p}(\tau_n)\}_{n=1}^N$. The distribution of $f_{x,p}(Z)$ is given by

$$\mathrm{Pr}_{x,\theta}[f_{x,p}(Z) = f_{x,p}(t)] = \sum_{f_{x,p}(\tau_n) = f_{x,p}(t)} \mathrm{Pr}_{x,\theta}[Z = \tau_n]$$
$$= \sum_{f_{x,p}(\tau_n) = f_{x,p}(t)} f_{x,\theta}(\tau_n), \tag{11}$$

where $f_{x,\theta}$ is given by equation (A-13). It is equivalent to use either $T(Z) = g_{x,p}(Z) = 1 - f_{x,p}(Z)$ or the rank $T(Z) = g_{x,p}^*(Z)$ of $g_{x,p}(Z)$ in $\{g_{x,p}(\tau_n)\}_{n=1}^N$, both of which have large values significant. Again, $g_{x,p}^* = f_{x,p}^-$.

In either case, the cdf of $T(Z)$ is

$$F_{x,\theta}(t) = \sum_{n=1}^{N} f_{x,\theta}(\tau_n) I_{[T(\tau_n),\infty)}(t), \quad \text{i.e.,}$$

$$= \sum_{n=1}^{N} f_{x,\theta}(\tau_n) I_{[g_{x,p}(\tau_n),\infty)}(t) \quad \text{or,}$$

$$= \sum_{n=1}^{N} f_{x,\theta}(\tau_n) I_{[g_{x,p}^*(\tau_n),\infty)}(t), \tag{12}$$

and the null distribution is given when $\theta = p$. The decision rule is to reject if $T(Z) > T_\alpha$, where the critical value is $T_\alpha = Q_{x,p}(1-\alpha_0)$. The power of the test under the alternative (x,θ) is $1-\beta = 1 - F_{x,\theta}(T_\alpha)$.

In practice there may be "degenerate" distributions, which arise when there is a single number coming out of X in some component. That is to say $n_i = 1$, $X_i = x_i$, and $p_i = 1$ for some i. If $L_i \neq x_i$, reject H_0. Otherwise, test the nondegenerate components as above.

Of course, the multinomial model subsumes the binomial model. With model output and data limited to $\{0,1\}$ it is certain that L lies in S, the sample space of the model X. This is not necessarily so for the multinomial model. For some i it may be that L_i does not lie in $\{x_{i,j}\}_{j=1}^{n_i}$. This event occurs with null probability zero, hence ranks higher than any vector the model produces. And so the null hypothesis is rejected here.

2.2   The Need for Sensitivity

Analysts have attempted to modify the OP test in order to eliminate the strict requirement that $L \in S$. Such modifications usually invoke some comparison of L_i with model output $x_{i,j}$, for example, through using histogram techniques or "binning close values together," which implies the construction of a distance measure on a space containing the model output X and data vector L. In other words, this introduces into the analysis a concept of distance between model LoF vectors. The OP test does not use the model output values, only their probabilities of occurrence. Model prediction values are used as an index to the probabilities that are being compared. So the actual data and model values have no part in the OP test and are only introduced in an effort to generalize the test and as an aid in understanding OP test results. The OP test statistic cannot provide any indication of the effect of these adjustments. However, it is possible to construct a test that incorporates model outcome values (in addition, of course, to their probabilities of occurrence) and enables quantification of such effects.

A similar difficulty can arise in the application of the OP test even when $L \in S$. Suppose the test results in rejection. The analyst may wonder why and ask if a particular component is responsible for the reject decision. The usual tack taken here is to examine the location of L among the values of X to gain insight about the behavior of $g_{x,p}^*(L)$ among $g_{x,p}^*(X)$. Again, the OP test can only provide information when the real (or altered) experiment results lie in the simulation sample space. The new test to be proposed can accommodate arbitrary incremental adjustments to the experiment data and allow for true sensitivity analysis.

This situation also arises in power consideration. A significance level α may be chosen to manipulate β against a certain alternative, seeking a low β against a "highly unlikely" outcome. The unlikeliness of an outcome is most likely to be judged by its distance from likely outcomes in terms of value and not probability.

Sometimes the OP test gives surprising results. Consider a multinomial example in one dimension with $x = (0, 0.15, 0.9)$ and $p = (0.8, 0.04, 0.16)$. Statistics are $g_{x,p}^*$ $= (1, 3, 2)$ with cdf values $(0.8, 1, 0.96)$. Set the significance level at $\alpha_o = 0.16$. Then $T_\alpha = 2$ and the reject region is $g_{x,p} \in \{3\}$, or $x \in \{0.15\}$. If the data is $L = 0.15$, the test rejects H_0. Note that the expected value of the model is $E\,X = 0 \times 0.8 + 0.15$ $\times 0.04 + 0.9 \times 0.16 = 0.15$. So the OP test asserts that the data does not come from the model distribution when the data is in fact the mean of the model distribution and is in the model sample space. Note that if the data is $L = 0.9$, the OP test does not reject H_0. This value of x is far removed from the mean of the distribution in the following sense. Consider a statistic that is the rank of distance from the mean. That is, $T^* = (2, 1, 3)$ with cdf values $(0.84, 0.04, 1)$. With $\alpha = 0.16$, the test based on this statistic has $T_\alpha^* = 2$ and so rejects H_0 when $L = 0.9$ but not when $L = 0.15$. On one hand, $L = 0.15$ is a rare event and $L = 0.9$ is more likely to occur. On the other hand, $L = 0.15$ lies nearer to a group of more likely outcomes than does $L = 0.9$.

It is beneficial to quantify the concept of distance that is already being used in conjunction with OP test analysis of LF experiments. This can be accomplished with the use of a statistic and hypothesis test based on a measure of central tendency in the model/simulation output value space. Such a test is now proposed.

3. The MX Test

The VL model with output X in the sample space $S = \otimes_{i=1}^{k} \{x_{i,j}\}_{j=1}^{n_i}$ is k-dimensional vector multinomial $M_k(1, x, p)$ with value parameter list x and

probability parameter list p. A test of the hypothesis

$$H_0 : \mathrm{E}\,L = \mathrm{E}\,X$$
$$H_1 : \mathrm{E}\,L \neq \mathrm{E}\,X \tag{13}$$

is conducted using the test statistic $T = M_{x,p}$ given by

$$M_{x,p}(Z) = \sum_{i=1}^{k} \frac{(Z_i - \mathrm{E}\,X_i)^2}{\mathrm{Var}\,X_i} = \sum_{i=1}^{k} \Delta_i (Z_i)^2$$

$$\text{where} \quad m_i = \mathrm{E}\,X_i, \quad s_i^2 = \mathrm{Var}\,X_i, \quad \text{and} \quad \Delta_i(t) = \frac{t - m_i}{s_i}. \tag{14}$$

This is a test of $\mathrm{E}\,L = \mathrm{E}\,X$ given that $\mathrm{Var}\,L = \mathrm{Var}\,X$ under H_0. It is not required that $L \in S$. The distribution of $M_{x,p}$ is, of course,

$$F(t) = \Pr[M_{x,p}(Z) \leqslant t]. \tag{15}$$

When $Z \sim M_k(1,z,\theta)$, this is

$$F_{z,\theta}(t) = \sum_{n=1}^{N} f_{z,\theta}(\tau_n) I_{[M_{x,p}(\tau_n),\infty)}(t), \tag{16}$$

where $Z \in \otimes_{i=1}^{k} \{z_{i,j}\}_{j=1}^{n_i} = \{\tau_n\}_{n=1}^{N}$ and $N = \prod_{i=1}^{k} n_i$, with $f_{z,\theta}$ given by equation (A-13). Note that the τ_n comprise an enumeration of the Z sample space, not the sample space of the simulation X. The distribution parameter (x,p) of X enters the calculation through the indicator of $[M_{x,p}(\tau_n),\infty)$; and the probability that $M_{x,p}(Z)$ assumes a particular value is

$$\Pr_{z,\theta}[M_{x,p}(Z) = M_{x,p}(\tau_n)] = \sum_{I} \Pr_{z,\theta}[Z = \tau_i]$$
$$= \sum_{I} f_{z,\theta}(\tau_i), M_{x,p}(\tau_i) = M_{x,p}(\tau_n)\}, \tag{17}$$

where $I = \{i \mid M_{x,p}(\tau_i) = M_{x,p}(\tau_n)\}$, which is a sum of probabilities from the distribution of Z. The null distribution of $M_{x,p}(Z)$ is given when $Z \sim M_k(1,x,p)$.

The statistic $M_{x,p}$ has an approximate chi-square distribution, $M_{x,p} \sim \chi_k^2$, under the null hypothesis. The distribution of $M_{x,p}$ under an alternative hypothesis is approximately noncentral chi-square, $M_{x,p} \sim \chi_{k,\delta}^2$, with the usual noncentrality parameter $\delta = \sum_{i=1}^{k}(\mathrm{E}\,Z_i - \mathrm{E}\,X_i)^2 / \mathrm{Var}\,X_i$.

This test, based on a multinomial distribution with a χ^2 approximation, will be referred to as the MX test.

4. Demonstration

For illustrative purposes, consider a test for the binomial model with parameter $p = (0.1, 0.2, 0.3, 0.4, 0.8)$ at the significance level $\alpha_o = 0.1$. Power is calculated against the specific alternative $\theta = (0.9, 0.2, 0.3, 0.6, 0.8)$. Distributions of the test statistics are shown under the null and alternative hypotheses, along with α, power, and values of the test statistics. Figure 1 shows the MX test statistic $M_{x,p}$ with its χ^2 approximation. Figure 2 shows the rank $M_{x,p}^*$ of the MX statistic. Figure 3 shows the OP test statistic $g_{x,p}$. Figure 4 shows the rank $g_{x,p}^*$ of the OP statistic. Correspondence of the null ranks is shown in Figure 5, along with the sizes and probabilities of the reject and accept regions. True significance levels are $0.025 + 0.066 = 0.091$ for the OP test and $0.066 + 0.018 = 0.084$ for the MX test. Note that $\Pr[\text{reject MX} \mid \text{reject OP}] = 0.066/0.091 = 0.73$ and that $\Pr[\text{reject OP} \mid \text{reject MX}] = 0.066/0.084 = 0.79$. So, these two different tests do not always yield the same decision. The probability that a random point passes one test and fails the other is $0.025 + 0.018 = 0.043$, and 3 points out of 32 lie in this region.

To assess the general power characteristics of these tests consider, for fixed dimension k, specifying a null distribution $(p_i)_{i=1}^k$ with uniform random components and testing this against a uniform random alternative $(q_i)_{i=1}^k$ for fixed α_o. This gives a random value for the Type II error probability β. Repeat this 2500 times to get empirical distributions, which are estimates of the distributions of β for a test of a random null against a random alternative. Estimated median power $1 - \beta$ is shown for various values of k in Table 1. The power of the OP is greater by as much as 2%, so there is little appreciable difference between the two tests.

5. Application

Data from a series of 14 LF tests will now be used to illustrate these statistical procedures by comparing the LF results to predictions generated by the MUVES/S2 simulation model. The LF LoF assessment data L and parameters for the LoF simulation distribution X are presented in Table 2 for MLoF (mobility LoF) and Table 3 for FLoF (firepower LoF). The parameters of the simulation model are LoF values x and their respective probabilities of occurrence p. Tables 4 and 5 include respectively for MLoF and FLoF the additional quantities involved in the calculation of the MX statistic detailed in equation (14).

First, consider the MLoF data. The five degenerate (single-valued) distributions of shots 3, 8, 9, 10, and 13 agree with their LF results, so we proceed to analyze the remaining set of size 9. Note that L_4, L_6, and L_{12} do not lie in the sample space of the simulation distribution. So the OP test is not applicable here.

However, one may argue that, in practice, the L_4 value of 0.26 is "close enough" to the value of 0.22 in the model distribution for that shot for the two to be considered equal. This justifies changing the value of L_4 to 0.22 for the purposes of conducting the OP test. Likewise, $L_6 = 1$ and $x_{6,4} = 0.96$ are "close", so we change L_6 to 0.96. Then, noting that $L_{12} = 0.04$ is close to $x_{12,2} = 0.02$, we declare that $L_{12} = 0.02$. The modified LF vector now lies in the simulation sample space, and the OP test can be applied. The resulting p-value is 0.19, insufficient evidence to reject discrepancy at the 10% significance level.

On the other hand, $L_{12} = 0.04$ is equally close to $x_{12,3} = 0.06$. So we might as well declare that $L_{12} = 0.06$. The OP test in this case has a p-value of 0.086, indicating disagreement between the model and the experiment at the 10% level.

So, for this particular example, the OP test outcome can be manipulated at will to conclude consistency or inconsistency. Also, the effects of the adjustments made in this process cannot be quantified.

Now consider application of the MX test to the MLoF data. Even though the LF results from shots 4, 6, and 12 do not lie in the simulation distribution they are considered "close enough" to justify use of the MX test. Unlike the application of the OP test in dealing with this discrepancy, no alteration of the data is necesary for the MX test.

The p-value for the exact MX test is 0.069, and the p-value of the χ^2 approximation to the MX distribution is 0.12. These levels most likely but marginally indicate that H_0 should be rejected.

Shot L_4 has the greatest $|\Delta|$ of 3, so this may be the shot most responsible for the rejection. The data here is $L_4 = 0.26$, to be compared with X_4 parameters of $x_4 = (0.02, 0.06, 0.1, 0.22)$ and $p_4 = (0.36, 0.38, 0.13, 0.13)$, which give a mean of 0.0716 and a standard deviation of 0.0631. Deleting L_4 from the data set and performing the MX test on the remaining set of size 8 yields an exact p-value of 0.78, and an approximate p-value of 0.74.

So the MX test has been used here to isolate the single shot that is responsible for possible discrepancy between the simulation predictions and LF results.

In the manner of a sensitivity analysis, the effects on the p-value of altering the data values L_4, L_6, and L_{12} are shown in Table 6. The other six nondegenerate shots retain their original data values, which did not need to be altered for application of the OP test; and sensitivity analysis is conducted on the full nondegenerate data

10

set of nine shots. The MX test enables quantification of the effects of altering data, which is not possible relying solely on the OP test. This is certainly desirable. Clearly, the value of L_4 is the main factor in the outcome of the test. Given that $L_{12} = 0.04$ is close enough to 0.02 or 0.06, and that $L_6 = 1$ is close enough to 0.96, we see that changing L_4 from 0.26 to 0.24 results in a determination of model consistency at the 10% level, independent of the value of $L_{12} \in \{0.02, 0.04, 0.06\}$ or of $L_6 \in \{0.96, 1\}$. Therefore, given that $L_4 = 0.26$ is close enough to 0.24, we conclude that the MLoF model and LF results are consistent.

Similarly, changing the value of L_4 from 0.26 to 0.22 raises the exact p-value to 0.22. A reasonable conclusion to draw from this exercise is that marginal rejection is solely caused by underestimation on the part of the simulation by an amount in the neighborhood of 0.02–0.04 for shot 4.

Therefore, the decision to accept or reject model consistency hinges on the practical significance of this quantifiable discrepancy.

A similar consideration applies to the FLoF data. The p-values are essentially zero, and this can be attributed to shots 5 and 6 in which L_5 and L_6 are both equal to 0.05. As can be seen in Table 7, reduction of these data to somewhere between 0.02 and 0.03 results in p-values exceeding 0.2. The test is highly sensitive to the values of L_5 and L_6, and the original conclusion of inconsistency results from the model underestimating the LF responses of 0.05 for shots L_5 and L_6 by an amount in the neighborhood of 0.02–0.03.

Again, the MX test has allowed for quantification of the modeling discrepancy responsible for rejection of consistency, which then must be interpreted in the light of practical significance.

In conclusion, the MX test can be used to perform sensitivity analyses in tests for the consistency of live-fire results and simulation predictions. The MX test can be used to identify the critical shots that cause decisions of inconsistency for a series of predictions. The MX test can also be used to quantify the inconsistency in those predictions in a manner not possible with the OP test.
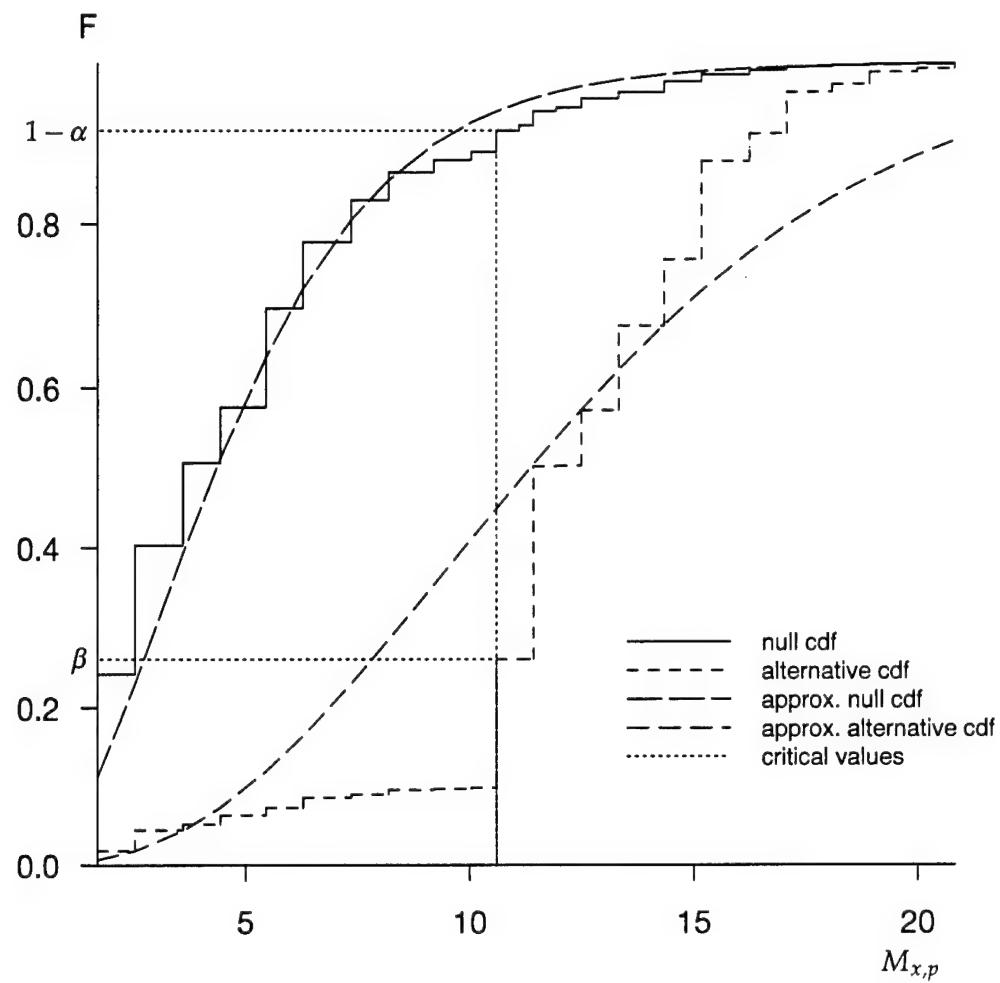
Figure 1. MX distribution.
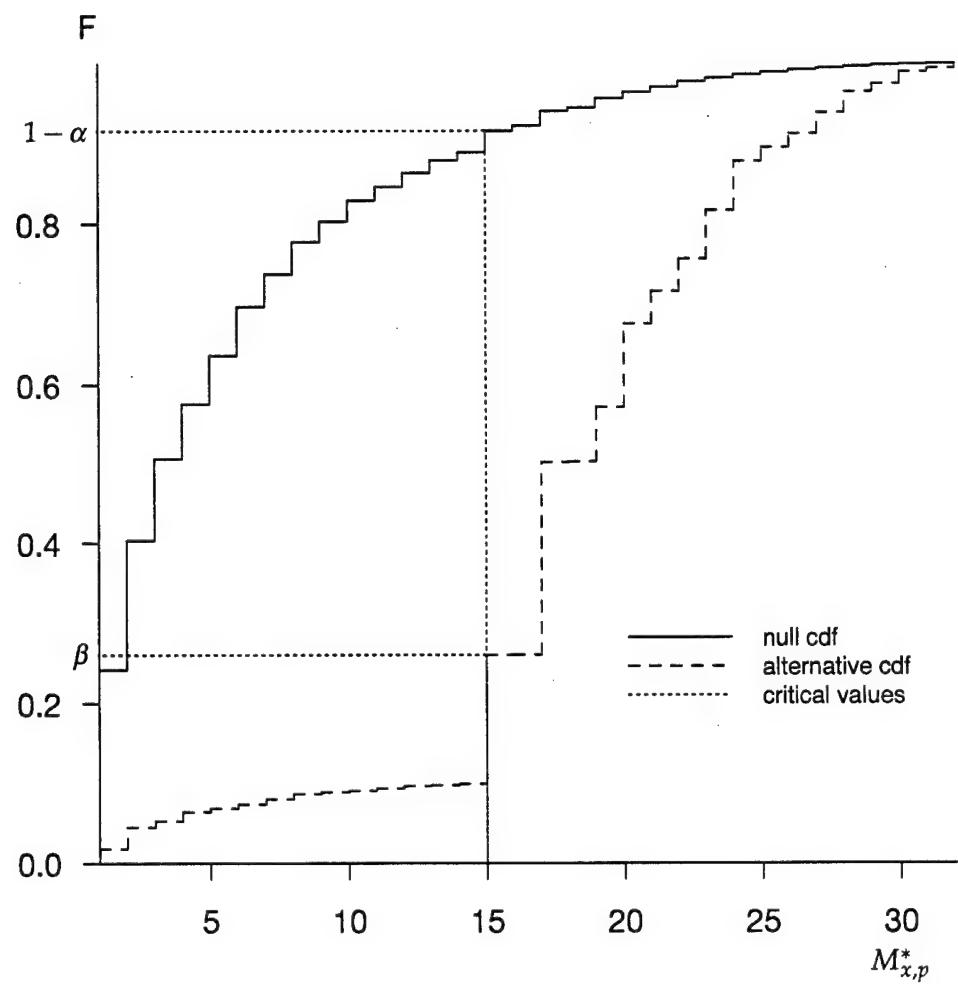
Figure 2. MX rank distribution.
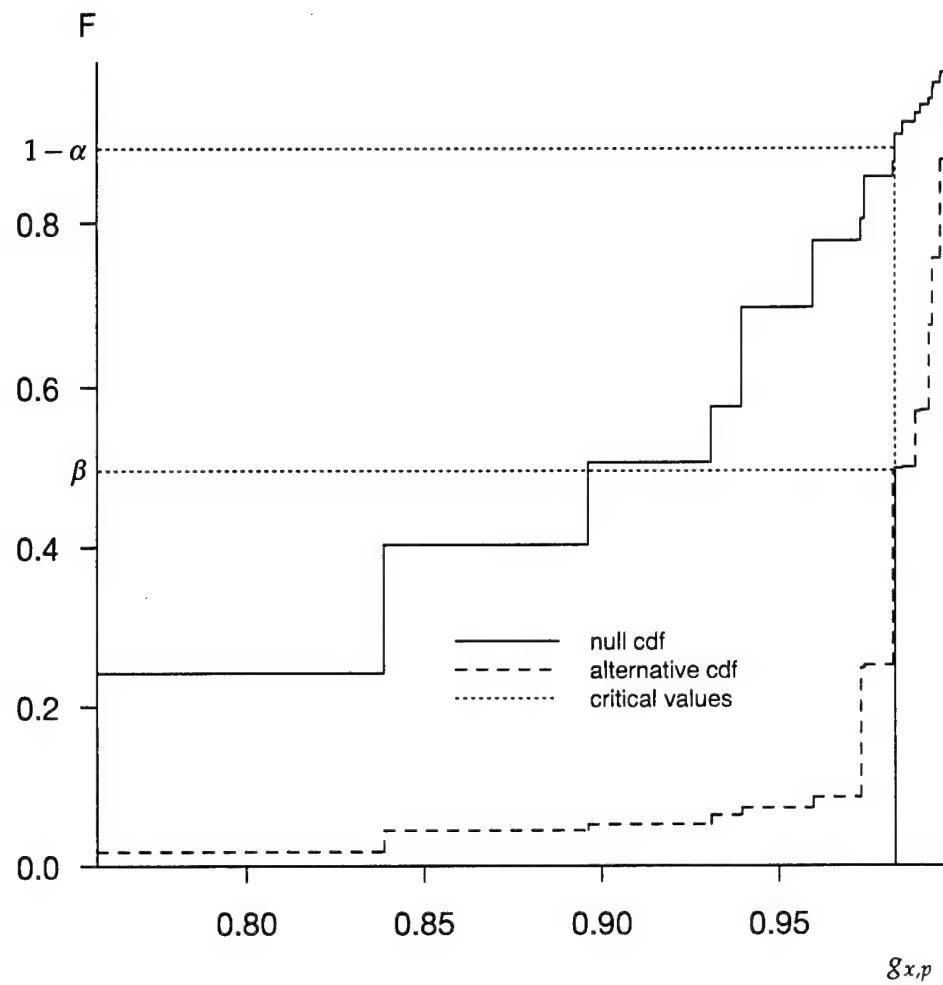
Figure 3. OP distribution.
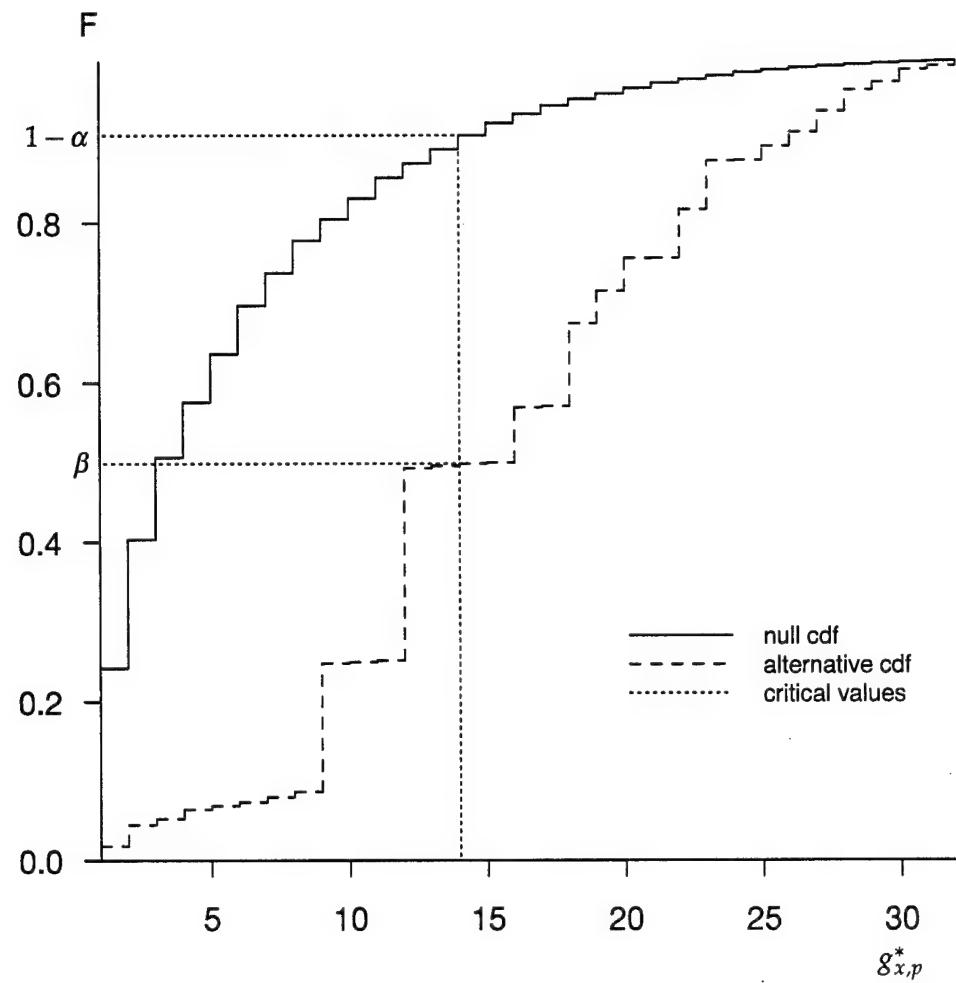
Figure 4. OP rank distribution.
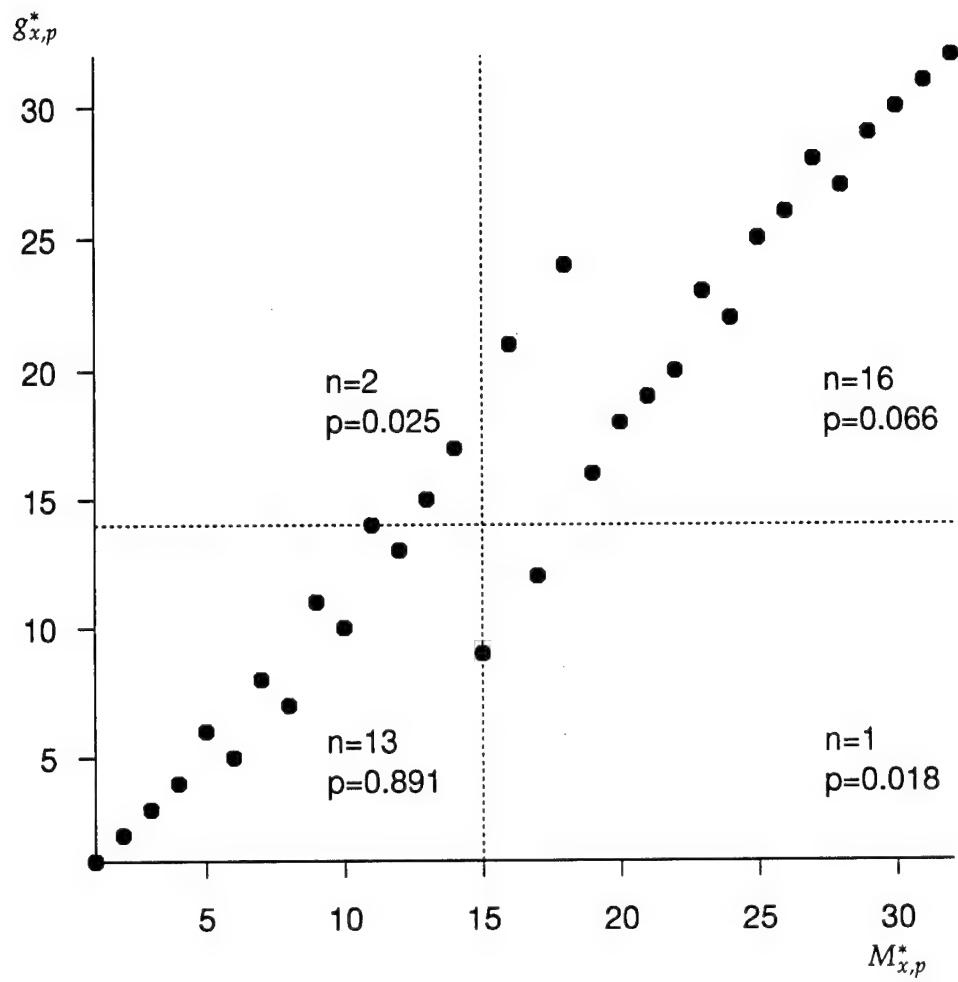
Figure 5. Rank-rank demonstration.

Table 1. Demonstration median power against random alternative.

k	OP	MX
5	0.62	0.61
6	0.69	0.69
7	0.76	0.74
8	0.78	0.77
9	0.84	0.82
10	0.85	0.84

Table 2. Basic MLoF data and model.

i	L_i	n_i	j	1	2	3	4	5
1	0	3	x	0	0.02	0.06		
			p	0.88	0.1	0.02		
2	1	2	x	0	1			
			p	0.32	0.68			
3	1	1	x	1				
			p	1				
4	0.26	4	x	0.02	0.06	0.1	0.22	
			p	0.36	0.38	0.13	0.13	
5	1	2	x	0	1			
			p	0.56	0.44			
6	1	4	x	0	0.02	0.8	0.96	
			p	0.21	0.29	0.46	0.04	
7	1	2	x	0	1			
			p	0.34	0.66			
8	0	1	x	0				
			p	1				
9	1	1	x	1				
			p	1				
10	0	1	x	0				
			p	1				
11	1	2	x	0	1			
			p	0.3	0.7			
12	0.04	5	x	0	0.02	0.06	0.1	0.26
			p	0.27	0.42	0.16	0.14	0.01
13	1	1	x	1				
			p	1				
14	1	2	x	0	1			
			p	0.18	0.82			

Table 3. Basic FLoF data and model.

i	L_i	n_i	j	1	2	3	4	5	6	7	8
1	0.98	2	x	0	0.98						
			p	0.06	0.94						
2	0.02	3	x	0	0.02	0.04					
			p	0.32	0.66	0.02					
3	1	1	x	1							
			p	1							
4	0.94	3	x	0.06	0.56	0.99					
			p	0.38	0.5	0.12					
5	0.05	2	x	0	0.02						
			p	0.56	0.44						
6	0.05	2	x	0	0.02						
			p	0.51	0.49						
7	0.02	2	x	0	0.02						
			p	0.34	0.66						
8	0	1	x	0							
			p	1							
9	0.02	1	x	0.02							
			p	1							
10	0	1	x	0							
			p	1							
11	1	2	x	0	1						
			p	0.3	0.7						
12	1	8	x	0	0.06	0.14	0.2	0.56	0.66	0.98	1
			p	0.12	0.17	0.06	0.02	0.05	0.01	0.43	0.14
13	1	1	x	1							
			p	1							
14	1	2	x	0	1						
			p	0.18	0.82						

Table 4. MLoF data and model.

i	L_i	m_i	s_i	Δ_i	n_i	j	1	2	3	4	5
1	0	0.0032	0.0101	−0.32	3	x	0	0.02	0.06		
						p	0.88	0.1	0.02		
2	1	0.68	0.466	0.69	2	x	0	1			
						p	0.32	0.68			
3	1	1	0	0	1	x	1				
						p	1				
4	0.26	0.0716	0.0631	3	4	x	0.02	0.06	0.1	0.22	
						p	0.36	0.38	0.13	0.13	
5	1	0.44	0.496	1.1	2	x	0	1			
						p	0.56	0.44			
6	1	0.412	0.402	1.5	4	x	0	0.02	0.8	0.96	
						p	0.21	0.29	0.46	0.04	
7	1	0.66	0.474	0.72	2	x	0	1			
						p	0.34	0.66			
8	0	0	0	0	1	x	0				
						p	1				
9	1	1	0	0	1	x	1				
						p	1				
10	0	0	0	0	1	x	0				
						p	1				
11	1	0.7	0.458	0.65	2	x	0	1			
						p	0.3	0.7			
12	0.04	0.0346	0.0403	0.13	5	x	0	0.02	0.06	0.1	0.26
						p	0.27	0.42	0.16	0.14	0.01
13	1	1	0	0	1	x	1				
						p	1				
14	1	0.82	0.384	0.47	2	x	0	1			
						p	0.18	0.82			

Table 5. FLoF data and model.

i	L_i	m_i	s_i	Δ_i	n_i	j	1	2	3	4	5	6	7	8
1	0.98	0.921	0.233	0.25	2	x	0	0.98						
						p	0.06	0.94						
2	0.02	0.014	0.01	0.60	3	x	0	0.02	0.04					
						p	0.32	0.66	0.02					
3	1	1	0	0	1	x	1							
						p	1							
4	0.94	0.422	0.313	1.70	3	x	0.06	0.56	0.99					
						p	0.38	0.5	0.12					
5	0.05	0.0088	0.00993	4.10	2	x	0	0.02						
						p	0.56	0.44						
6	0.05	0.0098	0.01	4.00	2	x	0	0.02						
						p	0.51	0.49						
7	0.02	0.0132	0.00947	0.72	2	x	0	0.02						
						p	0.34	0.66						
8	0	0	0	0	1	x	0							
						p	1							
9	0.02	0.02	0	0	1	x	0.02							
						p	1							
10	0	0	0	0	1	x	0							
						p	1							
11	1	0.7	0.458	0.65	2	x	0	1						
						p	0.3	0.7						
12	1	0.619	0.439	0.87	8	x	0	0.06	0.14	0.2	0.56	0.66	0.98	1
						p	0.12	0.17	0.06	0.02	0.05	0.01	0.43	0.14
13	1	1	0	0	1	x	1							
						p	1							
14	1	0.82	0.384	0.47	2	x	0	1						
						p	0.18	0.82						

Table 6. MLoF p-value sensitivity.

L_4	L_6	L_{12}	p-value		note
			MX	OP	
0.22	0.96	0.02	0.23	0.19	pass OP
0.22	0.96	0.04*	0.24	–	
0.22	0.96	0.06	0.21	0.086	fail OP
0.22	1*	0.02	0.21	–	
0.22	1*	0.04*	0.22	–	
0.22	1*	0.06	0.19	–	
0.24	0.96	0.02	0.14	–	
0.24	0.96	0.04*	0.14	–	
0.24	0.96	0.06	0.12	–	
0.24	1*	0.02	0.12	–	
0.24	1*	0.04*	0.13	–	
0.24	1*	0.06	0.11	–	
0.26*	0.96	0.02	0.074	–	
0.26*	0.96	0.04*	0.076	–	
0.26*	0.96	0.06	0.067	–	
0.26*	1*	0.02	0.067	–	
0.26*	**1***	**0.04***	**0.069**	–	true data vector
0.26*	1*	0.06	0.063	–	

"*" denotes actual data value

Table 7. FLoF p-value sensitivity.

L_4	L_5	L_6	p-value		note
			MX	OP	
0.94*	0.02	0.02	0.56	–	
0.94*	0.02	0.03	0.23	–	
0.94*	0.02	0.05*	0.037	–	
0.94*	0.03	0.02	0.21	–	
0.94*	0.03	0.03	0.087	–	
0.94*	0.03	0.05*	0.012	–	
0.94*	0.05*	0.02	0.028	–	
0.94*	0.05*	0.03	0.0091	–	
0.94*	**0.05***	**0.05***	**0.0000002**	–	true data vector
0.99	0.02	0.02	0.47	0.339	
0.99	0.02	0.03	0.19	–	
0.99	0.02	0.05*	0.031	–	
0.99	0.03	0.02	0.18	–	
0.99	0.03	0.03	0.078	–	
0.99	0.03	0.05*	0.0092	–	
0.99	0.05*	0.02	0.024	–	
0.99	0.05*	0.03	0.0073	–	
0.99	0.05*	0.05*	0	–	

"*" denotes actual data value

INTENTIONALLY LEFT BLANK.

Appendix. Background, Notation, and Terminology

The material in this section can be found in any standard text on mathematical statistics, such as Mood, Graybill, and Boes.[1] It is presented here for easy reference.

A.1   Some Useful Functions

The indicator function for the set S is

$$I_S(t) = \begin{cases} 1, & t \in S \\ 0, & t \notin S. \end{cases} \tag{A-1}$$

The cumulative distribution function (cdf) of a univariate random variable (rv) X is the left-continuous function

$$F(t) = \Pr[X \leqslant t], \quad t \in \mathbb{R} \tag{A-2}$$

and its right-continuous pseudo-inverse is the quantile function (qf)

$$Q(u) = \inf\{t : F(t) \geqslant u\}, \quad u \in [0,1]. \tag{A-3}$$

For continuous distributions $Q = F^{-1}$. For any rv, $Q(0) = -\infty$.

A discrete finite random variable X defined on the sample space $S = \{t_i\}_{i=1}^n$ has probability function $f(t) = \Pr[X = t]$ and cdf

$$F(t) = \Pr[X \leqslant t] = \sum_{i=1}^n \Pr[X = t_i]I_{[t_i,\infty)}(t) = \sum_{i=1}^n f(t_i)I_{[t_i,\infty)}(t). \tag{A-4}$$

The empirical cdf based on the i.i.d. data $\{X_i\}_{i=1}^n$ is the discrete cdf of n equiprobable points

$$F_n(t) = \frac{1}{n}\sum_{i=1}^n I_{[X_i,\infty)}(t) \tag{A-5}$$

and its inverse is the empirical qf $Q_n(u)$. For any (discrete, continuous, or mixed) rv

$$F(Q(u)) \geqslant u \quad \text{and} \quad Q(F(t)) \leqslant t. \tag{A-6}$$

---

[1]Mood, A. M., F. A. Graybill, and D. C. Boes. *Introduction to the Theory of Statistics*. 3rd edition, NY: McGraw-Hill, 1974.

A.2 Some Useful Random Variables

These particular random variables play a part in the procedures discussed and developed in this report. In general, $A \sim B$ means that the random variables A and B have the same distribution. So "$\sim$" can be read "has the same distribution as" or "is distributed as" or "is equal in distribution to."

A scalar Bernoulli(θ) or binomial $B(1,\theta)$ random variable Z is defined on $S = \{0,1\}$, and the parameter is $\theta = \Pr[Z = 1] = 1 - \Pr[Z = 0]$. So its probability function is

$$f_\theta(t) = \Pr_\theta[Z = t] = \theta^{I_{\{1\}}(t)}(1 - \theta)^{I_{\{0\}}(t)}, \quad t \in \{0,1\}, \tag{A-7}$$

and its moments are

$$\mathrm{E}\, Z = \theta \quad \text{and} \quad \mathrm{Var}\, Z = \theta(1 - \theta). \tag{A-8}$$

A sample of n independent $B(1,\theta)$ random variables has the usual binomial distribution $B(n,\theta)$. In the vector case, the binomial $B_k(1,\theta)$ random variable is $Z = (Z_i)_{i=1}^k$ with parameter $\theta = (\theta_i)_{i=1}^k$. Each component $Z_i \sim B(1,\theta_i)$, and they are independent. So Z is defined on $S = \{0,1\} \times \cdots \times \{0,1\} = \otimes_{i=1}^k \{0,1\} = \{0,1\}^k$. The probability function of Z is

$$f_\theta(t) = \Pr_\theta[Z = t] = \prod_{i=1}^k \Pr_\theta[Z_i = t_i] = \prod_{i=1}^k \theta_i^{I_{\{1\}}(t_i)}(1 - \theta_i)^{I_{\{0\}}(t_i)}, \tag{A-9}$$

where $t = (t_i)_{i=1}^k \in \{0,1\}^k$. Moments are given by

$$\mathrm{E}\, Z_i = \theta_i \quad \text{and} \quad \mathrm{Var}\, Z_i = \theta_i(1 - \theta_i). \tag{A-10}$$

A sample of n independent $B_k(1,\theta)$ random variables has the vector binomial distribution $B_k(n,\theta)$.

A scalar multinomial $M(1,z,\theta)$ random variable Z has parameter (z,θ) where $z = (z_j)_{j=1}^n$ and $\theta = (\theta_j)_{j=1}^n$ with $\sum_{j=1}^n \theta_j = 1$ and $0 < \theta_j \leqslant 1$ for each j. The distribution is given by $\Pr[Z = z_j] = \theta_j$, so the sample space is $S = \{z_j\}_{j=1}^n$. The probability function for Z is

$$f_{z,\theta}(t) = \Pr_{z,\theta}[Z = t] = \prod_{j=1}^n \theta_j^{I_{\{z_j\}}(t)}, \quad t \in \{z_j\}_{j=1}^n, \tag{A-11}$$

and its moments are $\mathrm{E}\, Z^m = \sum_{j=1}^n \theta_j z_j^m$, so

$$\mathrm{E}\, Z = \sum_{j=1}^n \theta_j z_j \quad \text{and} \quad \mathrm{Var}\, Z = \sum_{j=1}^n \theta_j z_j^2 - (\mathrm{E}\, Z)^2. \tag{A-12}$$

A sample of n independent $M(1,z,\theta)$ random variables has the usual multinomial distribution $M(n,z,\theta)$. The vector multinomial $M_k(1,z,\theta)$ random variable $Z = (Z_i)_{i=1}^{k}$ has parameter (z,θ), where $z = (z_i)_{i=1}^{k}$ and $\theta = (\theta_i)_{i=1}^{k}$. The ith component $Z_i \sim M(1,z_i,\theta_i)$, where $z_i = (z_{i,j})_{j=1}^{n_i}$ and $\theta_i = (\theta_{i,j})_{i=1}^{n_i}$. The sample space is $S = \otimes_{i=1}^{k}\{z_{i,j}\}_{j=1}^{n_i}$. The probability function is

$$f_{z,\theta}(t) = \Pr_{z,\theta}[Z = t] = \prod_{i=1}^{k}\Pr_{z,\theta}[Z_i = t_i] = \prod_{i=1}^{k}\prod_{j=1}^{n_i}\theta_{i,j}^{I_{\{z_{i,j}\}}(t_i)}, \qquad (A\text{-}13)$$

where $t = (t_i)_{i=1}^{k} \in \otimes_{i=1}^{k}\{z_{i,j}\}_{j=1}^{n_i}$, and moments are given by $\mathrm{E}\,Z_i^{m} = \sum_{j=1}^{n_i}\theta_{i,j}z_{i,j}^{m}$, so

$$\mathrm{E}\,Z_i = \sum_{j=1}^{n_i}\theta_{i,j}z_{i,j} \quad \text{and} \quad \mathrm{Var}\,Z_i = \sum_{j=1}^{n_i}\theta_{i,j}z_{i,j}^{2} - (\mathrm{E}\,Z_i)^2. \qquad (A\text{-}14)$$

In fact, $B_k(1,\theta) \sim M_k\big(1,(0,1),(1-\theta,\theta)\big)$. A sample of n independent $M_k(1,z,\theta)$ random variables has the vector multinomial distribution $M_k(n,z,\theta)$.

A.3   Hypothesis Testing

A hypothesis test is a procedure for determining the truth of some conjecture about the parameter of a probability distribution based on data, which is an observation from the distribution. Suppose the data U is a random variable defined on a sample space S according to a probability law with parameter θ, and that $T : S \to \mathbb{R}$ is a function so that $T(U)$ is a statistic. Consider the hypothesis test

$$H_0 : \theta = \gamma$$
$$H_1 : \theta \neq \gamma, \qquad (A\text{-}15)$$

which asserts that the correct value of θ is in fact γ. If small T indicates that $\theta = \gamma$ and large T indicates otherwise, then T is said to have large values significant. The decision rule is to reject H_0 if $T > T_\alpha$ where the critical level T_α is chosen so that the probability of an incorrect decision under the null hypothesis, or Type I error, is no greater than α_o. This requires that T_α is large enough that $\Pr_\gamma[T > T_\alpha] \leqslant \alpha_o$ but no larger. So choose

$$\begin{aligned}
T_\alpha &= \inf\{T : \Pr_\gamma[T > T_\alpha] \leqslant \alpha_o\}\\
&= \inf\{T : \Pr_\gamma[T \leqslant T_\alpha] \geqslant 1 - \alpha_o\}\\
&= \inf\{T : F_\gamma(t) \geqslant 1 - \alpha_o\}\\
&= Q_\gamma(1 - \alpha_o). \qquad (A\text{-}16)
\end{aligned}$$

Then the actual probability of Type I error (also called the size of the critical region or significance level of the test) is

$$\alpha = \Pr_\gamma[T > T_\alpha] = 1 - \Pr_\gamma[T \leqslant T_\alpha] = 1 - F_\gamma(T_\alpha) = 1 - F_\gamma(Q_\gamma(1 - \alpha_o))$$
$$\leqslant 1 - (1 - \alpha_o) = \alpha_o \tag{A-17}$$

as required. The p-value of the test, denoted p, is defined as

$$p = 1 - F_\gamma(T). \tag{A-18}$$

If $p < \alpha$, then $F_\gamma(T) > F_\gamma(T_\alpha)$ and $T > T_\alpha$. Thus, $p < \alpha$ implies that H_0 is to be rejected.

The Type II error probability β of making an incorrect decision under an alternative with the specific parameter value θ is

$$\beta = \Pr_\theta[T \leqslant T_\alpha] = F_\theta(T_\alpha), \tag{A-19}$$

and the probability of making the correct decision under θ is the power of the test

$$1 - \beta = \Pr_\theta[T > T_\alpha] = 1 - F_\theta(T_\alpha). \tag{A-20}$$

$(X_{(i)})_{i=1}^n$ denotes the order statistics of $(X_i)_{i=1}^n$, so that $X_{(i)} \leqslant X_{(i+1)}$. And X_n^* denotes rank, so $X_{(i)}^* = i$. Descending rank will be denoted by X_n^-, so $X_{(i)}^- = n + 1 - i$.

NO. OF COPIES	ORGANIZATION
2	DEFENSE TECHNICAL INFORMATION CENTER DTIC OCA 8725 JOHN J KINGMAN RD STE 0944 FT BELVOIR VA 22060-6218
1	HQDA DAMO FDT 400 ARMY PENTAGON WASHINGTON DC 20310-0460
1	OSD OUSD(A&T)/ODDR&E(R) DR R J TREW 3800 DEFENSE PENTAGON WASHINGTON DC 20301-3800
1	COMMANDING GENERAL US ARMY MATERIEL CMD AMCRDA TF 5001 EISENHOWER AVE ALEXANDRIA VA 22333-0001
1	INST FOR ADVNCD TCHNLGY THE UNIV OF TEXAS AT AUSTIN 3925 W BRAKER LN STE 400 AUSTIN TX 78759-5316
1	US MILITARY ACADEMY MATH SCI CTR EXCELLENCE MADN MATH THAYER HALL WEST POINT NY 10996-1786
1	DIRECTOR US ARMY RESEARCH LAB AMSRL D DR D SMITH 2800 POWDER MILL RD ADELPHI MD 20783-1197
1	DIRECTOR US ARMY RESEARCH LAB AMSRL CI AI R 2800 POWDER MILL RD ADELPHI MD 20783-1197

NO. OF COPIES	ORGANIZATION
3	DIRECTOR US ARMY RESEARCH LAB AMSRL CI LL 2800 POWDER MILL RD ADELPHI MD 20783-1197
3	DIRECTOR US ARMY RESEARCH LAB AMSRL CI IS T 2800 POWDER MILL RD ADELPHI MD 20783-1197

ABERDEEN PROVING GROUND

2	DIR USARL AMSRL CI LP (BLDG 305)

29

30

1	US ARMY TRADOC ANL CTR ATRC W A KEINTZ WSMR NM 88002-5502
1	US ARMY RESEARCH OFFICE 4300 S MIAMI BLVD RESEARCH TRIANGLE PARK NC 27709
1	US ARMY RESEARCH LAB AMSRL SL C HOPPER WSMR NM 88002-5513
1	US ARMY RESEARCH LAB AMSRL SL E J PALOMO WSMR NM 88002-5513
1	US ARMY RESEARCH LAB AMSRL SL EA R FLORES WSMR NM 88002-5513
1	US ARMY RESEARCH LAB AMSRL SL EI J NOWAK FT MONMOUTH NJ 07703-5601

ABERDEEN PROVING GROUND

1	SBCCOM RDEC AMSSB RTD J ZARZYCKI 5183 BLACKHAWK RD APG MD 21010-5424
1	US ARMY DEV TEST COM CSTE DTC TT T APG MD 21005-5055
1	US ARMY EVALUATION CENTER CSTE AEC W HUGHES 4120 SUSQUEHANNA AVE APG MD 21005-3013

ABERDEEN PROVING GROUND (CONT)

1	US ARMY EVALUATION CENTER CSTE AEC SV L DELATTRE 4120 SUSQUEHANNA AVE APG MD 21005-3013

12	DIR USARL
	AMSRL SL
	  DR WADE
	  J BEILFUSS
	AMSRL SL B
	  J FRANZ
	  L ROACH
	AMSRL SL BA
	  M RITONDO
	AMSRL SL BD
	  J MORRISSEY
	AMSRL SL BE
	  P TANENBAUM
	AMSRL SL BG
	  D BELY
	AMSRL SL BN
	  D FARENWALD
	AMSRL SL E
	  M STARKS
	AMSRL SL EC
	  E PANUSKA
	AMSRL SL EM
	  J FEENEY

NO. OF
COPIES   ORGANIZATION

ABERDEEN PROVING GROUND

1   DIR USAMSAA
        AMXSY EI
        R CHANDLER
        APG MD 21005

28   DIR USARL
        AMSRL SL BA
           G BRADLEY
           E DAVISSON
           R DIBELKA
           S JUARASCIO
           D LYNCH
           J SHORT
           A VOGT
           W WINNER
        AMSRL SL BD
           J COLLINS (5 CPS)
           P MERGLER
           L MOSS
           E SNYDER
        AMSRL SL BE
           R BOWERS
           M MAHAFFEY
           R SAUCIER
           R SHNIDMAN
        AMSRL SL BG
           M ENDERLEIN
           T MUEHL
           K MURRAY
           J PLOSKONKA
           S PRICE
           J ROBERTSON
           N THOMPSON
           L WILSON

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project(0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	February 2002	Final, October 2000–March 2001

4. TITLE AND SUBTITLE
Sensitivity Analysis in Testing the Consistency of Vulnerability Models

5. FUNDING NUMBERS
622618AH80

6. AUTHOR(S)
Joseph C. Collins

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
ATTN: AMSRL-SL-BD
Aberdeen Proving Ground, MD 21005-5068

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-2668

9. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Comparison of series of live-fire test results with stochastic simulation model output has been a long-standing problem. The usual approach, the order by probability (OP) test, allows neither for quantification of typical adjustments made in the implementation of the test nor for sensitivity analysis. A new test is presented here which may be used in conjuction with the OP test to provide these missing features.

14. SUBJECT TERMS
statistics, sensitivity analysis, live-fire testing

15. NUMBER OF PAGES
37

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18   298-102

INTENTIONALLY LEFT BLANK.